

Knihovna pro určení vzájemně podobných fotografií vhodného pro produkční provoz

Bc. Dobroslav Pelc



*** Nascanované zadání, strana 1 ***

*** Nascanované zadání, strana 2 ***

Prohlašuji, že

- beru na vědomí, že odevzdáním diplomové práce souhlasím se zveřejněním své práce podle zákona č. 111/1998 Sb. o vysokých školách a o změně a doplnění dalších zákonů (zákon o vysokých školách), ve znění pozdějších právních předpisů, bez ohledu na výsledek obhajoby;
- beru na vědomí, že diplomové práce bude uložena v elektronické podobě v univerzitním informačním systému dostupná k prezenčnímu nahlédnutí, že jeden výtisk diplomové práce bude uložen v příruční knihovně Fakulty aplikované informatiky. Univerzity Tomáše Bati ve Zlíně a jeden výtisk bude uložen u vedoucího práce;
- byl/a jsem seznámen/a s tím, že na moji diplomovou práci se plně vztahuje zákon č. 121/2000 Sb. o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon) ve znění pozdějších právních předpisů, zejm. § 35 odst. 3;
- beru na vědomí, že podle § 60 odst. 1 autorského zákona má UTB ve Zlíně právo na uzavření licenční smlouvy o užití školního díla v rozsahu § 12 odst. 4 autorského zákona;
- beru na vědomí, že podle § 60 odst. 2 a 3 autorského zákona mohu užít své dílo – diplomovou práci nebo poskytnout licenci k jejímu využití jen připouští-li tak licenční smlouva uzavřená mezi mnou a Univerzitou Tomáše Bati ve Zlíně s tím, že vyrovnání případného přiměřeného příspěvku na úhradu nákladů, které byly Univerzitou Tomáše Bati ve Zlíně na vytvoření díla vynaloženy (až do jejich skutečné výše) bude rovněž předmětem této licenční smlouvy;
- beru na vědomí, že pokud bylo k vypracování diplomové práce využito softwaru poskytnutého Univerzitou Tomáše Bati ve Zlíně nebo jinými subjekty pouze ke studijním a výzkumným účelům (tedy pouze k nekomerčnímu využití), nelze výsledky diplomové práce využít ke komerčním účelům;
- beru na vědomí, že pokud je výstupem diplomové práce jakýkoliv softwarový produkt, považují se za součást práce rovněž i zdrojové kódy, popř. soubory, ze kterých se projekt skládá. Neodevzdání této součásti může být důvodem k neobhájení práce.

Prohlašuji,

- že jsem na diplomové práci pracoval samostatně a použitou literaturu jsem citoval. V případě publikace výsledků budu uveden jako spoluautor.
- že odevzdaná verze diplomové práce a verze elektronická nahraná do IS/STAG jsou totožné.

Ve Zlíně

.....

podpis autora

ABSTRAKT

Cílem této práce je analýza možností pro určení vzájemně podobných fotografií. Na základě analýzy student vybere nejvhodnější návrh řešení pro potřeby reálného produkčního provozu. Výslednou komponentu realizuje formou distribuované služby.

Klíčová slova: Přehled klíčových slov

ABSTRACT

Text of the abstract

Keywords: Some keywords

Zde je místo pro případné poděkování, motto, úryvky knih, básní atp.

OBSAH

ÚVOD	9
I TEORETICKÁ ČÁST	9
1 KOEFICIENT PODOBNOSTI DVOU FOTOGRAFIÍ	11
1.1 ZMĚNA VELIKOSTI FOTOGRAFIE	11
1.2 PŘEVOD NA SVĚTLOSTNÍ MATICI	12
1.3 PŘEVOD DO VLNOVÉHO SPEKTRA	13
1.4 KŘÍŽOVÁ KORELACE (CROSS CORRELATION METHOD)	13
1.5 PŘEVOD ZPĚT Z VLNOVÉHO SPEKTRA	14
2 POUŽITÉ ANALYTICKÉ METODY	16
2.1 DIFERENČNÍ ANALÝZA	16
3 KLASIFIKACE ŘEŠENÝCH OBLASTÍ	17
3.1 KONSTRUKTIVNÍ ZMĚNY FOTOGRAFIE	17
3.1.1 Návrh řešení	17
3.1.2 Oblast zájmu	17
3.2 DESTRUKTIVNÍ ZMĚNY FOTOGRAFIE	17
3.2.1 Návrh řešení	17
3.2.2 Oblast zájmu	17
3.3 KOMBINACE KONSTRUKTIVNÍCH A DESTRUKTIVNÍCH ZMĚN	18
3.3.1 Návrh řešení	18
3.3.2 Oblast zájmu	18
3.4 VÝBĚR REPREZENTATIVNÍHO VZORKU	18
3.4.1 Návrh řešení	18
3.4.2 Oblast zájmu	18
3.5 PODNADPIS	18
II ANALYTICKÁ ČÁST	18
4 BRAINSTORMING	21
4.1 AKADEMICKÝ KRUH	21
4.2 PROFESNÍ KRUH	21
5 DIFERENČNÍ ANALÝZA (GAP ANALÝZA)	22
5.1 POPIS SOUČASNÉHO STAVU	22
5.2 POPIS CÍLOVÉHO STAVU	22
5.2.1 Nefunkční požadavky	22
5.3 ROZDÍLY	22

5.4	NÁVRH VARIANT K DOSAŽENÍ CÍLE	23
5.4.1	Výpočet koeficientů na CPU	23
5.4.2	Výpočet koeficientů na CPU s paralelizací na GPU	23
5.4.3	Výpočet koeficientů na PC farmě	23
5.5	ZHODNOCENÍ VARIANT	23
5.5.1	Výpočet koeficientů na CPU	24
5.5.2	Výpočet koeficientů na CPU s paralelizací na GPU	24
5.5.3	Výpočet koeficientů na PC farmě	24
6	BENCHMARKING	25
6.1	TESTOVACÍ PROSTŘEDÍ	25
6.1.1	Použitý HW	25
6.1.2	Použitý SW	25
6.2	VÝSLEDKY TESTOVÁNÍ	28
6.2.1	Výpočet koeficientů na CPU	28
6.2.2	Výpočet koeficientů na CPU s paralelizací na GPU	29
6.2.3	Výpočet koeficientů na PC farmě	30
III	PROJEKTOVÁ ČÁST	30
7	DISTRIBUOVANÁ SLUŽBA	32
7.1	FRONTA NEZPRACOVANÝCH OBRÁZKŮ	32
7.2	SERVLET PRO STAŽENÍ OBRÁZKŮ	32
7.3	ODESLÁNÍ VÝSLEDKŮ	32
8	KLIENT	33
	ZÁVĚR	34
	SEZNAM POUŽITÉ LITERATURY	35
	SEZNAM POUŽITÝCH SYMBOLŮ A ZKRATEK	36
	SEZNAM OBRÁZKŮ	37
	SEZNAM TABULEK	38
	SEZNAM PŘÍLOH	39

ÚVOD

První odstavec pod nadpisem se neodsazuje, ostatní ano (pouze první řádek, odsazení vertikální mezy odstavci je typické pro anglickou sazbu; czech babel toto respektuje, netřeba do textu přidávat jakékoliv explicitní formátování, viz ukázka sazby tohoto textu s následujícím odstavcem).

Formátování druhého odstavce. Text text text text text text text text text text text.

I. TEORETICKÁ ČÁST

1 Koeficient podobnosti dvou fotografií

Základní ukazatelem, který bude v této práci řešen je koeficient podobnosti (dále jen KoP). KoP leží na intervalu $< 0, 1 >$. Přičemž hodnoty blíží se 1 symbolizují podobné fotografie. Interval podobnosti byl na základě testovacích pokusů stanoven na $< 0, 007, 1)$. Hodnota 1 znamená duplicitní fotografii. Byla z intervalu vyloučena, jelikož jsou fotografie nejprve unifikovány pomocí otisku MD5 [1]. Výpočet KoP provedeme v šesti krocích, z čehož jsou čtyři kroky přípravné (optimalizační) a pouze dva kroky reálně ovlivňují výsledný KoP.

1. Změna velikosti fotografie
2. Převod na světlostní matici
3. Převod do vlnového spektra
4. Křížová korelace (Cross correlation method)
5. Převod zpět z vlnového spektra
6. Výpočet KoP z výsledné matice

Body 1 - 3 slouží jako přípravné a aplikují se na obě fotografie (vzorová a referenční). Do bodu 4 tedy vstupují dvě matice (pro každou fotografii jedna). Výstupem 4. bodu je již jen jedna matice, která je v bodě 6 vyhodnocena do výsledného KoP.

1.1 Změna velikosti fotografie

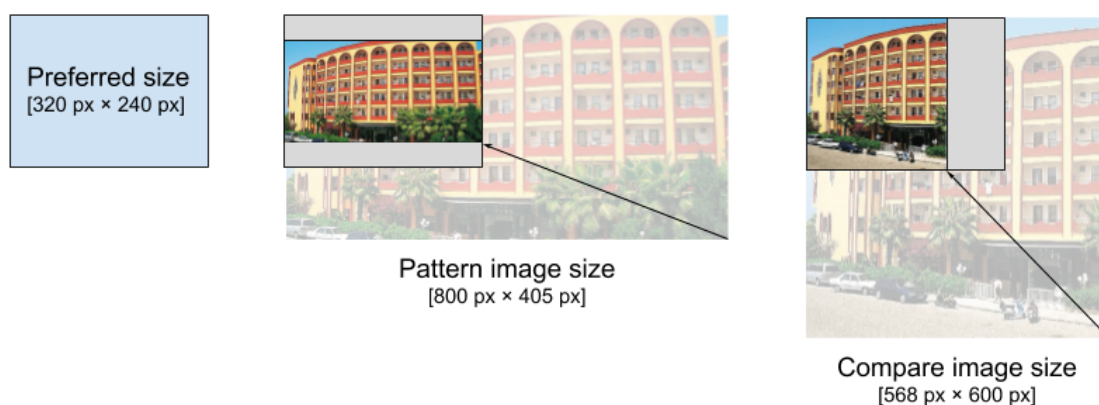
Jedná se o nezbytný přípravný krok, jehož cílem sjednotit u obou fotografií počet bodů a tím pádem také počet prvků v maticích, které vzniknou v následujícím kroku výpočtu KoP. Jako referenční velikost byla stanovena

- Šířka: 320 px,
- Výška: 240 px,

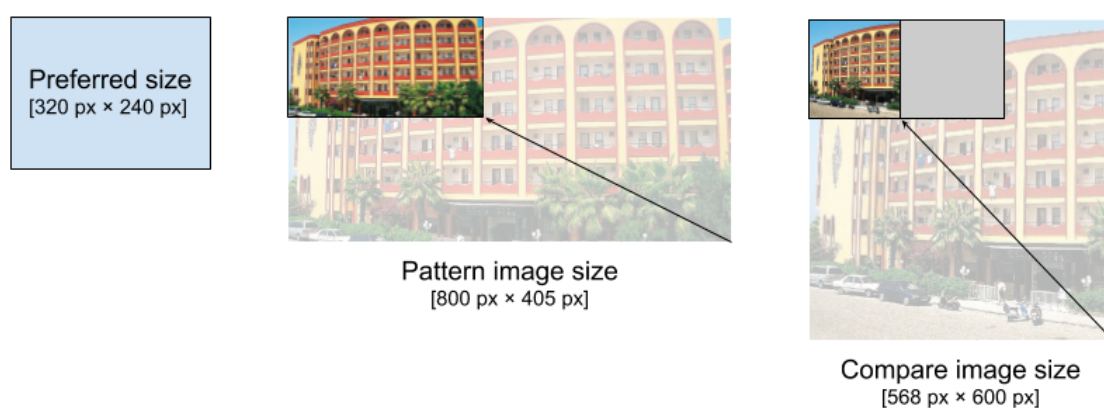
které vyšli jako nejvhodnější kompromis mezi relevancí výsledku a HW časem nutným k jeho zpracování.

Změna velikosti vzorové fotografie na $[320 \times ?]$ nebo $[? \times 240]$ se provádí v závislosti delší straně. Kratší strana je dopočítána podle původního poměru stran. Výsledný rozměr není doplněn nulami na plnou referenční velikost $[320 \times 240]$. Vzorová fotografie tedy určuje velikost, na kterou musíme upravit referenční fotografii. Pokud má referenční fotografie jiný poměr stran, bude doplněna nulami, aby nevznikaly prázdná místa.

Takto připravené fotografie nejsou při dalším zpracování komutativní, pokud mají vzájemně jiný poměr stran. V důsledku to znamená, že se musí porovnat fotografie v obou směrech (jak vzorová vůči referenční, tak referenční vůči vzorové). Můžete si to ukázat na příkladu (Obr. 1.1) a (Obr. 1.2).



Obr. 1.1 Statický scaling obrázků (je komutativní)



Obr. 1.2 Dynamický scaling obrázků (není komutativní)

Přesto, že je tento proces HW dražší, než jeho komutativní varianta, získáme díky tomu řádově lepší relevanci výsledků (zejména pokud je jedena ze dvou fotografií velmi nekvalitní, případně má nižší nativní rozlišení, než je referenční).

1.2 Převod na světlostní matici

Světlostní matice je fotografie co by rastrová bitmapa převedená do dvourozměrné matice (někdy též šedotónový obraz; v anglicky psaných textech k nalezení pod názvem brightness-matrix). Formálně je to dvourozměrná diskretní veličina, reprezentovaná maticí druhého řádu [2]. Každý bod původní fotografie - pixel [3] (dále jen px) je z RGB [4] hodnot převeden na hodnotu intenzity jasové funkce.

Původní px je reprezentován jako tří prvkové pole s hodnotami na intervalu $< 0, 255 >$.

- R => red,
- G => green,

- $B \Rightarrow \text{blue}$,

Výsledná hodnota intenzity jasové funkce p_x se spočítá jako střední hodnota z hodnot jednotlivých složek p_x (R, G a B). Jak napovídá interval možných hodnot, jeden p_x převedený na prvek světlostní matice zabírá v operační paměti 1 Byte (1 Byte = 8 bit \Rightarrow osmibitová barevná hloubka \Rightarrow 256 stupňů šedi). Na jednu plnou referenční fotografii je tedy potřeba 75 KB.

1.3 Převod do vlnového spektra

Převodem do vlnového spektra dosáhneme zajímavé výkonnosti optimalizace [5]. Kdybychom tento krok z celého procesu výpočtu KoP vynechali (stejně jako následně nezbytný převod zpět z vlnového spektra), museli bychom udělat křížovou korelaci v normálním spektru. Tzn. $O(n^4)$ operací, kde n je velikost strany čtvercové matice.

Pokud ale nejprve provedeme dopřednou Fourierovu transformaci (DFT) pro převod do vlnového spektra, a až následně křížovou korelaci, ušetříme jeden řád hodnot počtu operací. Budeme potřebovat $O(n^3 * \log n)$ operací. To je za předpokladu matice reprezentující 200×200 px fotografii rozdíl dvou řádů operací.

- Klasické spektrum: $200^4 = 1.6 * 10^9$ operací
- Vlnové spektrum: $200^3 * \log 200 = 6.1 * 10^7$ operací

1.4 Křížová korelace (Cross correlation method)

Korelace je nejdůležitější krok celého výpočtu KoP. Umožní pomocí jednoduchých matematických operací rozhodnout, zda jsou světlostní matice vzorové a referenční fotografie podobné. Míra podobnosti je vyjádřena hodnotami korelačních koeficientů sestavených do jedné matice (zatím ještě ve vlnovém spektru). Čím více se výsledné hodnoty blíží 1, tím více jsou fotografie podobné v daném bodě.

Celý proces není citlivý na konstruktivní změny (viz níže). Dokáže tedy podobnost vyhodnotit bez ohledu na změnu sytosti barev, kontrastu či jasu. Stejně tak výsledek není ovlivněn přidáním vodotisku nebo šumem. Přidání loga do fotografie již sice sníží KoP, ale pokud není logo přes polovinu fotografie, je stále bezpečně rozpoznána jako podobná či nikoliv.

Formálně je korelace zejména statistický pojem, který označuje vzájemný lineární vztah mezi znaky nebo veličinami. Míra korelace je daná korelačním koeficientem [6]. Zjednodušený postup výpočtu:

- K matici A se spočítá rozptyl σ_A a střední hodnota μ_A .
- K matici B se spočítá rovněž rozptyl σ_B a střední hodnota μ_B .

- Výslednou matici lze získat skalárním součinem A a B ponížených o jejich střední hodnoty.
- Výslednou matici je nutné na závěr vydělit součinem rozptylů obou matic.

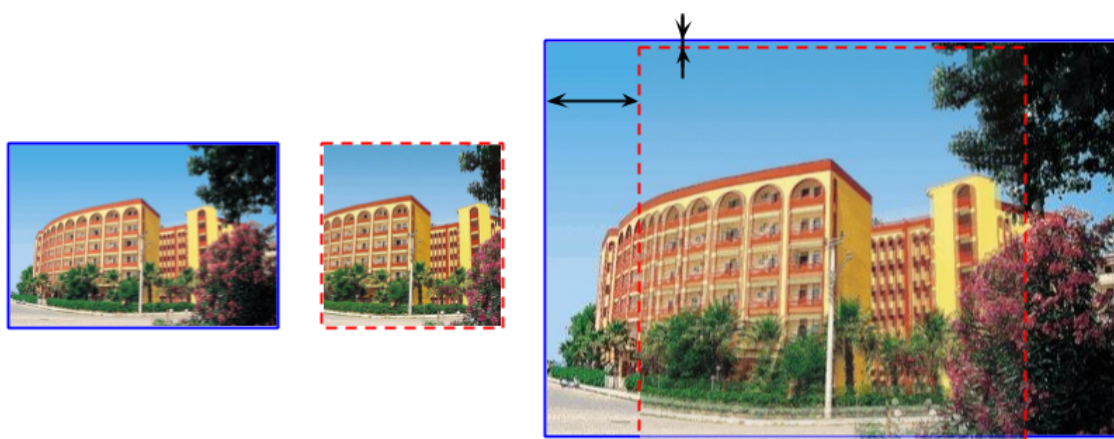
$$\rho_{A,B} = \frac{(A - \mu_A) \times (B - \mu_B)}{\sigma_A \sigma_B}$$

Za předpokladu, že objekty na porovnávaných fotografiích nejsou vůči sobě v prostoru posunuty, poskytuje korelační matice odpověď na otázku, zda jsou si dvě fotografie podobné. V praxi ale tento model příliš nenastává. Naopak je velmi časté, že porovnáváte např. výřez z fotografie oproti původní originální fotce, případně posunutou fotografii po libovolné z os.

Tuto problematiku řeší křížová korelace (Cross correlation method) [7]. Princip samotné korelace je stejný, pouze se opakuje s částečným posunem tak, aby pokryla všechny možné kombinace mezi dvěma fotografiemi. Příklad je asi nejlépe vidět na původních fotografiích, jak ukazuje (Obr. 1.3) a (Obr. 1.4).



Obr. 1.3 Příklad proložení dvou fotografií s použitím klasické korelace



Obr. 1.4 Příklad proložení dvou fotografií s použitím křížové korelace

1.5 Převod zpět z vlnového spektra

FFTW používá matlab i java

Křížová korelace ve vlnovém spektru - FT převede do vlnového spektra - pak proběhne křížová korelace - IFT pak převede zpět z vlnového spektra

Úspora převodem v normálním spektru křížová korelace je operací $O(n^4)$; n je jedna strana obrazové matice (čtvercové) ve vlnovém spektru křížová korelace je operací

Úspora je oproti původním

2 Použité analytické metody

2.1 Diferenční analýza

Diferenční analýzu (někdy též Gap analýza) navrhl Igor Ansoff. Skládá se z následujících kroků:

- Popis stávajícího stavu
- Stanovení cílů (popis cílového stavu)
- Určení rozdílu (mezery) mezi stávajícím a cílovým stavem
- Návrh variant dosažení cílového stavu (alternativní strategie)
- Zhodnocení variant a výběr nevhodnější z nich
- V případě potřeby se celý postup opakuje, dokud není dosaženo cílového stavu

3 Klasifikace řešených oblastí

TODO: navodit téma + uvést základní škálu

3.1 Konstruktivní změny fotografie

TODO: definovat konstruktivní změny

3.1.1 Návrh řešení

Křížová korelace zrychlená pomocí diskrétní Fourierovy transformace. TODO: rozvést

3.1.2 Oblast zájmu

TODO: uvést

Změny vlastností

- Změna barev; TODO: Popis změny barev
- Změna kontrastu; TODO: Popis změny kontrastu
- Změna jasu; TODO: Popis změny jasu

Změny obsahu

- Vodotisk
- Logo
- Šum

3.2 Destruktivní změny fotografie

TODO: definovat destruktivní změny

3.2.1 Návrh řešení

Výpočet Hausdorfovy vzdálenosti mezi konvexními polyedry, které reprezentující hrany v obrazu. TODO: rozvést

3.2.2 Oblast zájmu

- Změna komprese (rozmazaná fotka)
- Změny rozlišení
 - Ořez (v jedné nebo obou dimenzích)
 - Deformace (v jedné nebo obou dimenzích)

3.3 Kombinace konstruktivních a destruktivních změn

TODO: Obecně je porovnání problematické, rozvést.

3.3.1 Návrh řešení

Redukce fotografie na její prahovou velikost jako příprava na křížovou korelaci viz konstruktivní změny.

3.3.2 Oblast zájmu

- Asymetrická změna obou stran s čímkoliv
- Změna kvality v důsledku zhoršení komprese s čímkoliv
- Logo nebo vodotisk v kombinaci s předcházejícími

3.4 Výběr reprezentativního vzorku

Pro skupinu vzájemně si podobných fotografií vybereme nejvhodnějšího kandidáta, který bude následně ostatní fotografie zastupovat.

3.4.1 Návrh řešení

Pro tyto účely zavedeme koeficient podobnosti, který zjednodušeně určíme jako $VELIKOST \times OSTROST + JAS$. Výsledek bude na intervalu $<0, 1>$. Hodnoty blíží se nule mají nejnižší koeficient zastupitelnosti a naopak hodnoty blíží se 1 mají nejvyšší koeficient zastupitelnosti.

3.4.2 Oblast zájmu

Střední hodnota jasu Pouze zohlednění zda fotka není příliš jasná nebo příliš tmavá, jednoduchý algoritmus

Poměrný počet hran Test “rozmazanosti” fotky, konvoluce s vhodným (nutné testování) jádrem typu horní propust (s celkovým součtem 0) - tedy výsledek neutrální podklad z kterého “vystupují” hrany, velké množství hran \Leftrightarrow fotka není rozmazaná

Rozlišení fotografie Pouze klasická velikost fotky (větší \Leftrightarrow lepší).

3.5 Podnadpis

Tato sekce obsahuje ukázkou vložení tabulky (Tab. 3.1).

Tab. 3.1 Popisek tabulky

	1	2	3	4	5	Cena [Kč]
F	(jedna)	(dva)	(tři)	(čtyři)	(pět)	300

II. ANALYTICKÁ ČÁST

4 Brainstorming

Analytický metoda sloužící ke sběru myšlenek, námětů a případné zevrubné konstruktivní kritice dané problematiky. V rámci této práce byla použita v akademickém a profesním kruhu pro identifikaci základních ukazatelů pro další kroky analýzy.

4.1 Akademický kruh

Diskutovány zejména technické možnosti. Kde a jak lze vůbec porovnání fotografií provádět strojově. K další analýze vyly vyb

4.2 Profesní kruh

V kruhu s provozovatelem byly kladeny nejvyšší nároky na flexibilitu a propustnost celého řešení.

5 Diferenční analýza (GAP analýza)

5.1 Popis současného stavu

Je požadováno porovnání rastrových bitmap za účelem identifikace vzájemně podobných fotografií. Řešení je hledáno pro produkční provoz. Konzumentem cílového řešení je webový portál dovolena.cz, který má přibližně dva miliony fotografií. Průměrný počet přístupů k některé fotografii je přibližně 100 přístupů za sekundu. Jedná se především o hotely a jejich okolí. Webový portál slouží spíše jako datový konsolidátor. Nabídka portálu značně ovlivňuje cílové portfolio fotografií. Podle testovacích měření se za jeden týden obmění cca 10% fotografií z celkového množství. Jako testovací vzorek byl vybrán jeden nejmenovaný hotel a jeho 139 fotografií. Zákazník webového portálu vidí všechny fotografie v nesetříděné galerii. Některé fotografie jsou unikátní, ale většina si je velmi podobná. U některých dokonce nejsou lidským okem patrné rozdíly.

5.2 Popis cílového stavu

Konsolidované fotografie prezentované klientovi budou v maximální možné míře obsahovat unikátní fotografie. Vzájemně si podobné fotografie budou odfiltrovány a zůstane pouze jedna a to fotografie s nejvyšším indexem zastupitelnosti. Klient nebude čekat na zpracování podobnosti obrázků. Buď budou zpracované, pak klient uvidí jen unikáty, nebo ne a pak uvidí vše v původním stavu. V takovém případě se poměrově zvýší priorita na výpočet podobnosti fotek tohoto hotelu vůči ostatním ve frontě na výpočet. Cílové řešení musí být schopno operovat řádově s jednotkami milionů fotografií s týdenní fluktuací 15%.

5.2.1 Nefunkční požadavky

- Bezúdržbový systém
- Nevyžadující v průběhu času další financování
- Minimální vstupní investice
- Maximální kompatibilita s aktuálním HW

5.3 Rozdíly

- Nově vznikne nástroj pro určení koeficientu podobnosti dvou fotografií.
- Nově vznikne nástroj pro určení koeficientu zastupitelnosti vzájemně si podobných fotografií.
- Fotografie jednoho hotelu budou oindexovány a vnitřně škálovány do skupin pomocí koeficientů výše.
- Dojde k navýšení celkového počtu fotografií.
- Zvýší se také fluktuace fotografií (na očekávaných 15%).

5.4 Návrh variant k dosažení cíle

Pilířem celého řešení bude backendová strana cílového konzumenta. Limity a také jednotlivé možnosti pro realizaci jsou velmi omezeny nutností integrovat do současného řešení. Nejen z těchto důvodů má serverová strana spíše podpůrný charakter v projektu jako celku. Její význam je spíše v propojení všech jednotlivých komponent. Dojde tedy k modifikaci existujícího produkčního server-side prostředí. Nově zde poběží služba, která bude

- poskytovat zadání na určení koeficientů podobnosti a zastupitelnosti,
- poskytovat metadata nezbytná pro distribuci výpočtu,
- konzumovat výsledek distribuované operace,
- kompletovat zpracovaná data do cache vhodné pro silný organický provoz.

Naopak klientská strana je naprosto autonomní. Pro realizaci lze použít jak libovolnou platformu, tak libovolné technologie. Jediným technickým limitem je schopnost standardizovaným způsobem komunikovat se serverovou stranou.

5.4.1 Výpočet koeficientů na CPU

Základní myšlenka je využít nejdostupnější produkční HW a na zavedeném serveru spustit novou službu. Hlavní výhodou je dostupnost produkčního HW ve vlastním datacentru cílového konzumenta. Podstatnou nevýhodou fakt, že pro určení koeficientů výše není CPU ideální platforma.

5.4.2 Výpočet koeficientů na CPU s paralelizací na GPU

Základní myšlenka je osadit do produkčního serveru pracovní grafickou kartu a počítat podobnost fotografií na GPU.

5.4.3 Výpočet koeficientů na PC farmě

Základní myšlenka je využít HW osobních PC, kterých je v každé větší firmě požehnaně a neprovádět výpočet na jednom stroji, na každém dostupném stroji.

5.5 Zhodnocení variant

Zhodnocení vychází z benchmarkingu, který byl vyhodnocen jako pomocná analýza. Jako nejvýhodnější varianta pro realizaci vychází PC farma. Jako záložní řešení, lze využít produkční server s výpočtem na CPU. Pokud bude zvolena vhodný programovací jazyk, bude výsledné řešení přenositelné.

5.5.1 Výpočet koeficientů na CPU

- Za necelých 6 dní spočítá týdenní přírůstek.
- Za zbylou dobu z týdenního cyklu dále vypočítá přibližně 3,7% z celkového objemu sto milionu koeficientů.
- Pro plné vypočítání všech koeficientů potřebuje dalších cca 27 týdnů.
- Rezerva pro další růst (navýšení celkového počtu fotografií) je cca 18%.

5.5.2 Výpočet koeficientů na CPU s paralelizací na GPU

- Za necelých 8 hodin spočítá týdenní přírůstek.
- Za část ze zbylé doby týdenního cyklu dále vypočítá 100% z celkového objemu sto milionu koeficientů.
- Pro plné vypočítání všech koeficientů potřebuje celkem 2 dny.
- Rezerva pro další růst (navýšení celkového počtu fotografií) je cca 300%.

5.5.3 Výpočet koeficientů na PC farmě

Výsledek jednoho kancelářského PC je zanedbatelný a nemůže se rovnat předchozím variantám. Vezmeme-li v úvahu, že těchto strojů je k dispozici 16 hodin denně více než 300 kusů, začíná to již vypadat v číslech jinak.

- Za 2,5 hodiny spočítá týdenní přírůstek (za předpokladu 300 aktivních PC).
- Za část ze zbylé doby týdenního cyklu dále vypočítá 100% z celkového objemu sto milionu koeficientů.
- Pro plné vypočítání všech koeficientů potřebuje celkem 16 hodin, což je v tomto případě 1 den.
- Rezerva pro další růst (navýšení celkového počtu fotografií) je cca 700%.

Samozřejmě s tímto nestandardním krokem je spojená také revize infrastruktury, především kvalita a šířka pásma sítě, centrální správa PC aj, které je nutné v tuto chvíli zanedbat.

6 Benchmarking

Jednotlivé kandidáty pro realizaci klientské části (rozvedené v GAP analýze) podrobíme výkonnostním a srovnávacím testům. Základní předpoklady:

- Řádově je celkem potřeba jednorázově odbavit 100 000 000 výpočtů obou koeficientů.
- Na týdenní bázi následně odbavovat denní přírůstky (řádově 15 000 000 výpočtů obou koeficientů).
- Na denní bázi je to tedy něco přes 2 000 000 výpočtů obou koeficientů.

6.1 Testovací prostředí

6.1.1 Použitý HW

Simulace produkčního serveru Při volbě HW k testování byla zvolena pracovní stanice Lenovo, která je svoji sestavou velmi blízko produkčnímu serveru. Vzhledem k tomu, že disponuje starší generací CPU, bylo do ní osazeno i odpovídající GPU, aby bylo možné výsledky aproximovat na reální produkční HW.

- Lenovo ThinkStation D20
- Operační systém Windows 10 pro 64 bit
- CPU Intel(R) Xeon(R) X5670 @ 2.93 GHz (2×cpu, 6×core, 12×thread)
- RAM 32 GB DDR3 (1066)
- GPU NVIDIA GeForce GTX 460
- HDD SAS 10.000 ot

Simulace klasického PC Alternativní varianta řešení je využít v rozumné míře osobní PC. Cílový konzument řešení disponuje několika více než 300 ks kancelářských PC. Jako testovací vzorek byl vybrán zástupce drtivé většiny těchto PC.

- Dell optiplex 780
- Operační systém Windows 7 pro 32 bit
- CPU Intel Core 2 Duo E7500 2,93 GHz
- RAM 4 GB DDR3
- HDD SATA 250 GB

6.1.2 Použitý SW

V úvodní fázi projektu probíhal vývoj i testování algoritmů v Mathlabu. Mathlab vnitřně používá pro zpracování knihovnu FFTW [8]. Právě tato knihovna má dostupnou nativní implementaci v CMake [12]. Dále má pro spoustu programovacích jazyků připravený wrapper.

CPU

```

close all; clc; clear;
method = 'nearest';
type = 'single';
n = 35; %podle počtu obrázků...
%obrázky
im_gpu = cell(1, n);
h = zeros(1, n, type);
w = zeros(1, n, type);
s0 = 'hotel_images\aa';
%nahrávání obrázků
tic
for i = 1:n
    if i < 10
        s = [s0,'0',int2str(i),'.tif'];
    else
        s = [s0,int2str(i),'.tif'];
    end;
    im = single(imread(s)); %nahrání obrázku a převod
    [n1,n2,n3] = size(im);
    h(i) = n1;
    w(i) = n2;
    im_gpu{i} = im; %nahrání obrázku
end;
disp('Nahrávání všech obrázků:');
toc
%převádění obrázků na jasové matice
for i=1:n
    im_gpu{i} = sum(im_gpu{i}, 3) ./ 3; %převod na jasovou matici
end;
disp('Převod všech obrázků na jasovou matici:');
toc
log_cpu = zeros(n, n, type);
for i=1:n
    %ft prvního obrázku
    ft1 = fft2(im_gpu{i});
    ft1Norm = ft1 ./ abs(ft1);
    for j=1:n
        if i==j
            continue;
        end;
        scale = min(h(i)/h(j), w(i)/w(j));
        h2 = round(scale * h(j));
        h2 = min(h2, h(i));
        w2 = round(scale * w(j));
        w2 = min(w2, w(i));
        im2Sc = imresize(im_gpu{j}, [h2, w2], method);
        %doplnění druhého obrázku nulami na jednotnou velikost
        im2 = zeros(h(i), w(i), type);
        im2(1:h2, 1:w2) = im2Sc;
        %ft druhého obrázku
        ft2 = fft2(im2);
        ft2Norm = ft2 ./ abs(ft2);
        %křížová korelace ve vlnovém spektru
        ccW = ft1Norm .* conj(ft2Norm);
        %převod do obyčejného souřadnicového systému
        ccC = real(ifft2(ccW));
        %hledání maximálního korelačního koeficientu
        ccCMax = max(ccC);
        ccCMax = max(ccCMax);
        log_cpu(i,j) = ccCMax;
    end
end

```

```

        disp(['obr1: ',int2str(i),';  obr2: ',int2str(j)]);
    end;
end;
disp(['Cyklus ',int2str(n),' obrázků:']);
time = toc
clearvars -except log_cpu;

```

maximální koeficient ccCMax

GPU

```

%method = '_nearest';
%method = '_biquadr';
method = '_bicubic';
type = 'single'; %změnit zde a na ř. 32
n = 35;
path = 'kernels\ImgScale\ImgScale\kernel_';
kernel = parallel.gpu.CUDAKernel([path, type, method, '.ptx'], [path, type, method]);
%obrázky
im_gpu = cell(1,n);
h = zeros(1, n, type);
w = zeros(1, n, type);
s0 = 'hotel_images\aa';
%nahrávání obrázků na grafiku
tic
for i = 1:n
    if i < 10
        s = [s0,'0',int2str(i),'.tif'];
    else
        s = [s0,int2str(i),'.tif'];
    end
    im = single(imread(s)); %nahrání obrázku a převod
    [n1,n2,n3] = size(im);
    h(i) = n1;
    w(i) = n2;
    im_gpu{i} = gpuArray(im); %nahrání obrázku na grafiku
end
disp('Nahrávání všech obrázků do paměti grafiky:');
toc
%převádění obrázků na jasové matice
for i=1:n
    im_gpu{i} = sum(im_gpu{i}, 3) ./ 3; %převod na jasovou matici
end;
disp('Převod všech obrázků na jasovou matici:');
toc
log = zeros(n, n, type, 'gpuArray');
%cyklus
for i=1:n
    %ft prvního obrázku
    ft1 = fft2(im_gpu{i});
    ft1Norm = ft1 ./ abs(ft1);
    for j=1:n
        if i==j
            continue;
        end;
        scale = (min(h(i)/h(j), w(i)/w(j)));
        h2 = round(scale * h(j));
        h2 = min(h2, h(i));
        w2 = round(scale * w(j));
        w2 = min(w2, w(i));
    end
end

```

```

scale_h = h(j)/h2;
scale_w = w(j)/w2;
%im2Sc = imresize(im_gpu{j}, scale);
im2Sc = zeros(h2, w2, type, 'gpuArray');
kernel.GridSize = w2;
kernel.ThreadBlockSize = h2;
im2Sc = feval(kernel, im2Sc, im_gpu{j}, int32(h(j)), int32(w(j)), scale_h, scale_w);
%doplnění druhého obrázku nulami na jednotnou velikost
im2 = zeros(h(i), w(i), type, 'gpuArray');
im2(1:h2, 1:w2) = im2Sc;
%ft druhého obrázku
ft2 = fft2(im2);
ft2Norm = ft2 ./ abs(ft2);
%křížová korelace ve frekvenčním spektru
ccW = ft1Norm .* conj(ft2Norm);
%převod do obyčejného souřadnicového systému
ccC = real(iff2(ccW));
%hledání maximálního korelačního koeficientu
ccCMax = max(ccC);
ccCMax = max(ccCMax);
log(i,j) = ccCMax;
disp(['obr1: ',int2str(i),'; obr2: ',int2str(j)]);
end
end
disp(['Cyklus ',int2str(n),' obrázků:']);
log_gpu = gather(log);
toc
clearvars -except log_gpu;

```

6.2 Výsledky testování

Byly realizovány dva testovací scénáře.

- Výpočet čistě na CPU s maximálním využitím HW.
- Výpočet na CPU s využitím GPU pro paralelizaci vybraných procesů.

6.2.1 Výpočet koeficientů na CPU

Výpočet na CPU byl spuštěn v pěti vláknech. Z počátku vypadal průběh velmi slibně. Bohužel při delším testu se výpočet začal značně propadat. Později se ukázalo, že hlavním důvodem je odložené uvolňování RAM. Jinými slovy dokud byla další volná RAM, výpočet jel velmi rychle. S potřebou uvolnit RAM se výpočet řádově snížil.

- Krátkodobý test
 - doba: 5 minut
 - počet opakování testu: 10
 - průměrně za sekundu: 205 výpočtů obou koeficientů
- Dlouhodobý test
 - doba: 6 hodin
 - počet opakování testu: 10
 - průměrně za sekundu: 31 výpočtů obou koeficientů

- Propad o 85%

Na obrázku níže (Obr. 6.1) je výstup z profilování výpočtu na CPU. Obsahuje poměrově zvýrazněné dlouho trvající operace vůči zbytku procesu. V rámci profilování bylo zpracováno deset tisíc iterací a doby jednotlivých částí zprůměrovány.



Obr. 6.1 Vizualizace procesu výpočtu koeficientů na CPU

Celková zátěž alokovaného HW nebyla příliš slavná. Nepodařilo se efektivně využít CPU, které dosahovalo průměrné zátěže 30% zejména kvůli neustálé realokaci operační paměti, na kterou čekalo zkrátka vše. Úvodní myšlenka využít hrubou sílu produkčního HW se ukázala jako velmi špatná.

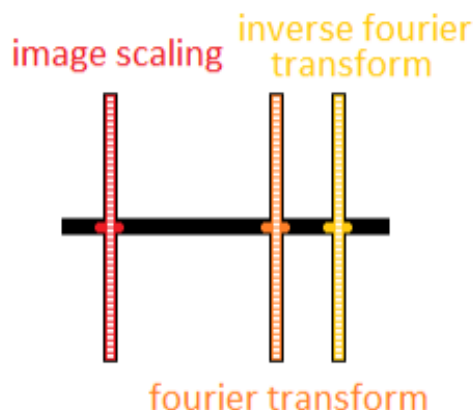
6.2.2 Výpočet koeficientů na CPU s paralelizací na GPU

Výpočet byl spuštěn v jednom vlákne na CPU. Vhodné operace pro GPU jsou delegovány pro paralelní zpracování. Nejprve bylo na GPU paralelizován pouze úvodní scaling obrázků. To nemělo příliš valný dopad na výsledky. Následně byly paralelizovány na GPU také prováděné transformace obrázků. Právě tento krok měl zásadní vliv na zrychlení celé operace. Je zde opět patrný lehký propad krátkodobého testu oproti dlouhodobému.

- Krátkodobý test
 - doba: 5 minut
 - počet opakování testu: 10
 - průměrně za sekundu: 617 výpočtů obou koeficientů
- Dlouhodobý test
 - doba: 6 hodin
 - počet opakování testu: 10
 - průměrně za sekundu: 559 výpočtů obou koeficientů
- Propad o 10%

Na obrázku níže (Obr. 6.2) je výstup z profilování výpočtu na CPU s využitím GPU pro paralelizaci dlouhotrvajících operací na CPU. Obsahuje poměrově zvýrazněné dlouho trvající operace vůči zbytku procesu. V rámci profilování bylo zpracováno deset tisíc iterací a doby jednotlivých částí zprůměrovány.

Zátěž na CPU dosahuje v průměru 5%. Naopak GPU dosahuje zátěže cca 75%. Otázkou zůstává, jak dlouho je schopná GPU pracovat pod tímto permanentním zatížením.



Obr. 6.2 Vizualizace procesu výpočtu koeficientů na CPU s paralelizací na GPU

6.2.3 Výpočet koeficientů na PC farmě

Výpočet byl spuštěn v jednom pracovním vlákne s omezenou možností alokace RAM na 1GB. Jedná se v podstatě o alternativu výpočtu obou koeficientů čistě na CPU. Stejně tak tomu odpovídal i celkový průběh zpracování, který téměř stejný, pouze s nižší dotací vypočítaných dvojic koeficientů. Lze tedy přejít rovnou na výsledky, jelikož samotné zpracování nic nového nepřineslo.

- Krátkodobý test
 - doba: 5 minut
 - počet opakování testu: 10
 - průměrně za sekundu: 10 výpočtů obou koeficientů
- Dlouhodobý test
 - doba: 6 hodin
 - počet opakování testu: 10
 - průměrně za sekundu: 6 výpočtů obou koeficientů
- Propad o 40%

Zátěž na CPU je konstantní. Jedno jádro ze dvou jede permanentně na 100%. Celková zátěž CPU tedy 50%. Vzhledem k tomu, že stoj již netrpěl na realokaci paměti, je v případě PC úzkým hrdlem opravdu CPU. Ostatní sledované metriky (vytížení HDD, síťový provoz, pracovní teplota) se příliš nevychýlili z normálu.

III. PROJEKTOVÁ ČÁST

7 Distribuovaná služba

7.1 Fronta nezpracovaných obrázků

7.2 Servlet pro stažení obrázků

7.3 Odeslání výsledků

8 Klient

ZÁVĚR

Text závěru

f

SEZNAM POUŽITÉ LITERATURY

- [1] MD5. In: *Wikipedia: the free encyclopedia* [online]. San Francisco (CA): Wikimedia Foundation, 2001- [cit. 2018-05-08]. Dostupné z: <https://en.wikipedia.org/wiki/MD5>
- [2] *Computer vision* [online]. [cit. 2018-05-08]. Dostupný z: http://midas.uamt.feec.vutbr.cz/ZVS/Exercise02/content_cz.php.
- [3] Pixel. In: *Wikipedia: the free encyclopedia* [online]. San Francisco (CA): Wikimedia Foundation, 2001- [cit. 2018-05-08]. Dostupné z: <https://cs.wikipedia.org/wiki/Pixel>
- [4] RGB. In: *Wikipedia: the free encyclopedia* [online]. San Francisco (CA): Wikimedia Foundation, 2001- [cit. 2018-05-08]. Dostupné z: <https://cs.wikipedia.org/wiki/RGB>
- [5] *IEEE Transactions on Computers* [online]. 1972, **C-21**(2), 179–186 [cit. 2018-05-08]. DOI: 10.1109/TC.1972.5008923. ISSN 0018-9340. Dostupné z: <http://ieeexplore.ieee.org/document/5008923/>
- [6] Korelace. In: *Wikipedia: the free encyclopedia* [online]. San Francisco (CA): Wikimedia Foundation, 2001- [cit. 2018-05-09]. Dostupné z: <https://cs.wikipedia.org/wiki/Korelace>
- [7] Cross-correlation. In: *Wikipedia: the free encyclopedia* [online]. San Francisco (CA): Wikimedia Foundation, 2001- [cit. 2018-05-09]. Dostupné z: <https://en.wikipedia.org/wiki/Cross-correlation>
- [8] *FFTW* [online]. [cit. 2018-05-08]. Dostupný z: <http://www.fftw.org>.
- [9] The Design and Implementation of FFTW3. In: *Proceedings of the IEEE* [online]. 2005, **93**(2), s. 216–231 [cit. 2018-05-08]. DOI: 10.1109/JPROC.2004.840301. ISSN 0018-9219. Dostupné z: <http://ieeexplore.ieee.org/document/1386650>
- [10] *IEEE Transactions on Image Processing* [online]. **5**(8), 1266–1271 [cit. 2018-05-08]. DOI: 10.1109/83.506761. ISSN 10577149. Dostupné z: <http://ieeexplore.ieee.org/document/506761/>
- [11] *Shared Scientific Toolbox in Java* [online]. [cit. 2018-05-08]. Dostupné z: <http://freshmeat.sourceforge.net/projects/shared>
- [12] *CMake* [online]. [cit. 2018-05-08]. Dostupný z: <https://cmake.org>.

SEZNAM POUŽITÝCH SYMBOLŮ A ZKRATEK

KoP	Koeficient podobnosti dvou fotografií
MD5	Message-Digest algorithm
px	picture element (obrazový prvek)
DFT	Discrete Fourier Transform (Diskrétní Fourierova transformace)
FFTW	Fastest Fourier Transform in the West
dddd	test

SEZNAM OBRÁZKŮ

Obr. 1.1	Statický scaling obrázků (je komutativní)	12
Obr. 1.2	Dynamický scaling obrázků (není komutativní)	12
Obr. 1.3	Příklad proložení dvou fotografií s použitím klasické korelace	14
Obr. 1.4	Příklad proložení dvou fotografií s použitím křížové korelace	14
Obr. 6.1	Vizualizace procesu výpočtu koeficientů na CPU	29
Obr. 6.2	Vizualizace procesu výpočtu koeficientů na CPU s paralelizací na GPU	30

SEZNAM TABULEK

Tab. 3.1	Popisek tabulky	19
----------	---------------------------	----

SEZNAM PŘÍLOH

P I.	Název přílohy
------	---------------

PŘÍLOHA P I. NÁZEV PŘÍLOHY

Obsah přílohy